

Stefan Weinzierl, Michael Vorländer

Room Acoustical Parameters as Predictors of Room Acoustical Impression: What Do We Know and What Would We Like to Know?

Journal article | Accepted manuscript (Postprint)

This version is available at <https://doi.org/10.14279/depositonce-9010>



This is a post-peer-review, pre-copyedit version of an article published in *Acoustics Australia*. The final authenticated version is available online at: <http://dx.doi.org/10.1007/s40857-015-0007-6>

Weinzierl, S., & Vorländer, M. (2015). Room Acoustical Parameters as Predictors of Room Acoustical Impression: What Do We Know and What Would We Like to Know? *Acoustics Australia*, 43(1), 41–48.
<https://doi.org/10.1007/s40857-015-0007-6>

Terms of Use

Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

WISSEN IM ZENTRUM
UNIVERSITÄTSBIBLIOTHEK

Technische
Universität
Berlin

Room Acoustical Parameters as Predictors of Room Acoustical Impression: What Do We Know and What Would We Like to Know?

Stefan Weinzierl · Michael Vorländer

Abstract Room acoustical parameters are audio features, usually extracted from monaural or binaural measurements of room acoustical environments, and used to predict different aspects of the ‘room acoustical impression’. The paper takes a closer look at the nature of this perceptual construct and at different approaches to develop a psychological measuring instrument for the multidimensional perceptual profile of room acoustical environments. Even after several decades of research, there is no satisfactory solution available for this purpose. The reasons for this lie in methodological deficits with respect to test development and item analysis as well as in a much too small sample of stimuli used in previous studies. Prospects for progress are opened up by state-of-the-art technologies for room acoustical simulation and auralization, which may be used to provide a large and representative sample of room acoustical environments as well as an authentic presentation in experimental studies. The fundamental perceptual components delivered by this approach will, most likely, not be predictable by traditional room acoustical parameters, but require advanced measurement techniques based on spherical arrays of transducers for both source and receiver characteristics, as well as new auditory models for feature extraction. The physical and psychological aspects of the problem are, in any case, inextricably linked with each other.

Keywords Room acoustics · Measurement and simulation · Room acoustical perception

1 Rooms as Technical and Perceptual Objects

In trying to capture the perceptual qualities of rooms, as a basis for room acoustical design or evaluation, one has to deal with the fact that rooms cannot be perceived as such, but only through their effect on the presented signal, the sound source, and the receiver involved. As with any other transmission system, its properties can only be studied in response to a given type of excitation. Assuming the room as a linear and time-invariant acoustical system, the output $Y(\omega)$ as the result of an excitation with the input signal $X(\omega)$ is given as

$$Y(\omega) = H(\omega) \cdot X(\omega) \quad (1)$$

with $H(\omega)$ as the (complex) transfer function of the room. If the measurement is contaminated with added noise $N(\omega)$, as it is always present in real environments, the transfer function can then be estimated by spectral division, yielding

$$\hat{H}(\omega) = \frac{Y(\omega) + N(\omega)}{X(\omega)} = H(\omega) + \frac{N(\omega)}{X(\omega)} \quad (2)$$

It should be noted that $H(\omega)$ includes the properties of source and receiver, in particular its directivities $G_S(\vartheta, \varphi)$ and $G_R(\vartheta, \varphi)$, which can be considered as weighting functions emphasizing certain propagation paths between source and receiver. The problem of directivities will be taken up again in Sect. 5. For the moment, it shall only be considered that $H(\omega)$ includes the specific directivities of source and receiver.

To analyse the behaviour of the system $H(\omega)$, we can choose basically any input signal, as long as it contains energy at all frequencies ω . Otherwise, the rightmost term in (2)

S. Weinzierl (corresponding author)
Audio Communication Group, TU Berlin, Berlin, Germany
e-mail: stefan.weinzierl@tu-berlin.de

M. Vorländer
Institute for Technical Acoustics, RWTH Aachen University,
Aachen, Germany
e-mail: mvo@akustik.rwth-aachen.de

will go to infinity in the presence of noise, and the spectral division will deliver no information about the system at ω . In practice, input signals are selected so that their spectral energy resembles that of the noise floor, so that the difference between the estimated transfer function $\hat{H}(\omega)$ and the “true” transfer function $H(\omega)$, given by the rightmost term in (2), tends to be frequency independent. Spectrally colored sine sweeps turned out to be a good and versatile choice for acoustical measurements in this regard [1].

By inverse Fourier transformation, the measured transfer function $H(\omega)$ can be translated into an impulse response $h(t)$, giving an easy-to-interpret representation in the time domain. Room acoustical parameters such as reverberation time (RT) or clarity (C_{80}) can be regarded as audio features extracted from $h(t)$.

Considering the process described in (1) from a psychological point of view, we have the same functional interaction, with the perceived signal $Y(\omega)$ depending on the input signal $X(\omega)$ (the audio content) and the room transfer function $H(\omega)$, including its dependence on the directivities of source and receiver and added noise. The essential difference between a technical and perceptual analysis is that the listener will not be able to carry out the deconvolution denoted in (2), which eliminates the influence of the source signal, with the same precision as this can be done analytically or numerically. The listener will always be confronted with the received signal as a whole, and his assessment of the latter, no matter in which form it is collected, will always be related to all its components, including the characteristics of the presented content.

One can, of course, draw the attention of the listener towards features of the received signal which are likely to be influenced by room acoustical properties. One may ask, for example, for the perceived ‘reverberance’ or for the degree of ‘envelopment’ produced by the presented musical signal, or for the degree of ‘intelligibility’ of speech, on the assumption that these qualities are primarily influenced by the acoustical reverberation rather than the signal itself. Thus, the listener is encouraged to carry out a kind of “perceptual deconvolution”, similar to the operation in (2), in order to separate the properties of excitation and transmission system. It is, however, difficult to predict to what extent this can be successfully performed. It is known, for example, that the early part of a reverberation tail tends to merge with the direct sound perceptually. It tends to increase the loudness of the sound source rather than being attributed to the spatial response. On the other hand, many musical instruments have their own decay phase, and with decay times of up to 3 s for violins on the open string [2], it is similarly difficult to separate the influence of source and room. From a gestalt psychological point of view, to realize this “perceptual deconvolution”, the listener has to construct two different auditory objects by grouping the perceived sound and attributing it to either “source” or “room”,

i.e. to construct a “source stream” and a “spatial stream” in the terminology of auditory stream segregation [3]. Since room reflections have a spatial and temporal origin different from the sound source, one would expect trained listeners to acquire the ability to assign them to the “spatial environment” as a separate auditory object within their environmental perceptual field. From a linguistic point of view, one could even argue that the fact that room acoustical experts quite naturally talk about ‘reverberance’ or ‘envelopment’ and perform relative assessments of it reflects spatial concepts and mental models of space acquired by professional practice and expertise [4]. This construction of the auditory objects “source” and “room” does, however, not necessarily reflect the corresponding physical components and, again, strongly depends on the presented content itself.

Aiming at a perceptual assessment of room acoustical environments, one can implement two different methodological strategies. The first is, to present identical signals to listeners in different acoustical environments. Employing partial correlation techniques, only the variance within impression ratings uniquely attributable to the factor “room” is then used as a dependent variable. This can be imagined as a kind of “post hoc deconvolution” of the influence of the presented content on the listeners ratings, allowing the possibility to evaluate the transmission system $H(\omega)$ alone. This was attempted by the majority of psychological studies as summarized in Sect. 2, and forms the basis for all technical analyses and the development of room acoustical parameters based on the measured room impulse response $h(t)$. The second strategy is to let the contribution of the presented content remain in the dependent variables, and to assess the receiver signal “as it is” in the real, functional context. The technical equivalent of this approach was adopted by some more recent investigations, aiming at features extracted from the receiver signal or a representative selection of typical content, respectively, directly, with auditory models for room acoustical perception as part of the algorithm applied [5].

2 Perceptual Properties and Room Acoustical Parameters: State of the Art

Throughout the first half of the twentieth century, the investigation of perceptual properties of room acoustical environments was mainly focused on the identification of preferred values for the reverberation time and its frequency dependence. First experiments were conducted already in 1902 by Sabine, who invited a number of musical experts to judge the acoustic quality of piano instruction rooms. He asked the test subjects to listen to piano music while seat cushions were successively added to the rooms in order to reduce their reverberation time, and observed that the listeners judged all rooms to be acoustically optimal if the reverberation time was within

a rather narrow range of tolerance [6]. Similar experiments were carried out by Bagenal with musicians as test subjects, who were asked to assess the reverberance and the effect of the room on the sound of their own musical instrument, while different materials were introduced into the room [7]. Other studies tried to interpolate the reverberation times of existing concert halls which were generally recognized for their superior acoustics, in order to find target values for acoustical planning [8–10]. In 1926, a first standard with guidelines for the reverberation time of rooms of different size was issued by the American Bureau of Standards [11].

After 1950, an increasing awareness can be observed, that an optimal reverberation time alone is no guarantee for a successful room acoustical design, and that ‘reverberance’ should not form the only criterion for the perceptual assessment of halls. Somerville and Gilford defined a glossary of 14 acoustic terms, which were “commonly used to describe the subjective qualities of a concert hall or studio” [12]. A similar list of 18 attributes was proposed by Beranek in his landmark book on *Music, Acoustics and Architecture*, along with relations between these perceptual qualities and physical properties of the hall, based on his own intuition and experience [13]. With 16 attributes selected from Beranek’s list, [14] conducted experiments in order to find underlying perceptual concepts by identifying latent variables, on which the ratings of the 16 attributes could be based, many of which turned out to be highly correlated among each other. By using factor analysis and multidimensional scaling (MDS), they arrived at different solutions with 4–6 independent factors.¹ While Hawkes and Douglas used their questionnaire in different British Concert Halls (with different musical programs and performers) and in the Royal Festival Hall in London, with the newly installed “Assisted Resonance” system in different technical settings, [15] used dummy head recordings of the Berlin Philharmonic Orchestra in six different halls, in order to vary acoustical stimuli experimentally and capture the assessment of subjects on a semantic differential with 19 different (German) attributes. Their factor analysis of room acoustic impression ratings delivered three latent variables, explaining 89% of the total variance. Considering the loadings of the original attributes on these variables, they were interpreted as ‘strength and extension of the sound source’, and ‘definition’ and ‘timbre’ of the overall sound [16]. By analyzing the bivariate correlations of the factor scores with 14 room acoustical parameters extracted from monaural impulse responses measured at each recording position, they tried to identify the best technical predictor for each of the above mentioned factor. While the strength factor G was highly correlated with the first factor ($r \approx 0.8$),

the predictive power of the single room acoustical parameters under test for the factors two and three were considered unsatisfactory, with explained variances of about 50 % and below.

The attributes used to describe perceptual qualities of the room acoustical impression in all studies mentioned above were always defined by the investigators themselves, based on their theoretical or practical experience with room acoustic design. In contrast, several studies appeared after 1990, aiming at an empirically substantiated approach in order to identify a vocabulary which can be assumed to be consistently used and interpreted by different subjects involved, as well as complete and yet as non-redundant as possible. Such studies, including a qualitative part for the verbal elicitation of the terminology and a quantitative part for the statistical analysis of the generated terms, were focused both on the evaluation of spatial audio reproduction systems [17] and on the perception of natural acoustical environments [18]. Using stimuli provided by impulse response measurements in eight different concert halls, encoded in Ambisonics B-Format and reproduced by a 14-channel loudspeaker system, Lokki et al. generated a vocabulary of 60 attributes, which were reliably used by 17 subjects involved, and which were elicited with an approach called individual vocabulary profiling. Based on the individual ratings of these attributes, the authors identified three principal components explaining 67 % of the total variance. Instead of a direct interpretation, these latent variables were considered in relation to clusters of attributes which were grouped independently by agglomerative hierarchical clustering. A group of attributes interpreted as ‘proximity’ descriptors was identified as crucial for the preference of the room acoustical environments presented.

As opposed to the individual elicitation of attributes a recent study dedicated to develop a psychological measuring instrument for the qualities of *simulated* acoustical environments used a focus group of experts in virtual acoustics and spatial audio technology in order to develop a consensus vocabulary of 48 attributes, as the result of a series of moderated roundtable discussions [19].

3 Research Gaps and Methodological Constraints

Recapitulating the fact that rooms cannot be perceived acoustically as such but only as a medium shaping the properties of the presented auditory content, it becomes obvious that the perceived phenomenal properties of this medium will depend partly on the properties of the presented content, the properties of the sound source involved, and the personal experience and preference of the listener addressed. An assessment of room acoustical environments on a one-dimensional ‘quality’ or ‘preference’ scale will thus only be possible if one narrows the view to a very limited choice of content, sources and listeners, such as 19th century sym-

¹ Although no details of the statistical approach are given in the paper, it can be assumed that Principal Component Analysis (PCA) was done, yielding *independent* factors.

phonic repertoire performed by standard-size symphonic orchestras for a homogeneous group of listeners with very specific expectations. Even then, one could ask from an artistic point of view, whether performance venues and their acoustical conditions could also serve to emphasize the variety of possible musical interpretations rather than being designed with regard to a technical standard or “optimum”. To expand upon this thought, however, goes beyond the focus of this contribution.

As soon as one leaves the narrow focus described above, the properties and the demands on rooms for music and speech can only be expressed as a multidimensional profile of perceptual features of room acoustical environments. With regard to the current state of the art there is, however, no satisfying measuring instrument for the concept of “room acoustical impression”. In view of the numerous efforts in this direction summarized above—what is missing yet?

First, none of the measuring instruments mentioned above, usually appearing in the form of semantic differentials consisting of between 10 and 20 individual attributes, is based on strong empirical evidence. All of them were defined ad hoc theoretically by the respective authors, and even if their professional experience in the field is taken into account, there are no objective indications of the quality of the resulting catalogue with respect to the standard criteria for psychometric tests. In the context of room acoustical impression, one would at least expect an analysis of

- external validity (How strong are measurements related to the acoustical properties of the rooms? In how far are ratings influenced by the raters’ experience, expertise, personal preference, and by the properties of the sources and the content involved?),
- item and construct reliability (Are the selected attributes as well as the latent variables of room acoustical perception interpreted consistently across time and individuals?),
- item discrimination (How well do single attributes serve to distinguish between different rooms?), which can be considered as an indicator of completeness of the item catalogue, and
- item difficulty (How well do single attributes and their scaling fit to the presented range of room acoustical properties?).

These criteria are part of standard item analyses for psychometric measurements and tests and may be performed post hoc, provided that enough measurements are available. This is, however, a crucial point: In order to determine the influence of personal traits and preferences and to have a representative sample of listeners with differing degree of expertise, a sufficient number of rating subjects is required. Second, in order to estimate the quality of the psychometric measuring instrument itself, sufficient variance and representativeness

within the presented pool of stimuli is required, i.e. a sufficient number of room acoustical environments has to be presented to the listeners involved as test subjects. Different criteria have been developed to quantify a minimum sample size, depending on the type of analysis to be performed. To mention only one criterion: A general identification of latent variables of room acoustical perception, i.e. a stable factor analytic solution of the measured data, which is valid beyond the specific sample of rooms used in the test, cannot be expected for a sample size of $n < 60$, i.e. less than 60 different room acoustical environments presented in the listening test. Even in a favorable case, i.e. with a good fit of attributes and factors (high communality), a sample of $N = 100$ is strongly recommended [20]. Comparing these requirements with the sample sizes used in the above mentioned room acoustical studies, with typically 6 [15] or 8 rooms [18], it becomes obvious that neither the dimension of the perceptual space, i.e. the number of latent factor variables, nor the structure and interpretation of the adopted factor solution can be reliably determined. The same deficit has been identified by [21] for the problem of room acoustical perception by musicians on stage.

Since the usefulness of room acoustical parameters is related to the extent to which they can predict single qualities of room acoustical impression, a good measuring instrument for this psychological construct is essential, if new parameters shall be developed for a focused room acoustical design. It will be hardly feasible to produce the large number of stimuli required for the experimental development of such a measuring instrument based on recordings in real environments alone. Thus, the progress in room acoustical simulation and auralization could be exploited to generate a sample of sufficient size. At the same time, only a *standard* instrument, e.g. a consensus terminology of room acoustical attributes, will allow for a comparison of different room acoustical evaluations and a meta-analysis of the results for methodological purposes.

4 Room Acoustical Measurements and the Reliability of Room Acoustical Parameters

In spite of a standardized framework [22,23], room acoustical measurements are affected by different sources of uncertainties, which limit the reliability to which room acoustical parameters can be determined in practice. These uncertainties refer to the definition of sources, receivers, and to the measurement process itself.

4.1 Measurement Algorithms

Various methods can be applied to obtain room impulse responses $h(t)$, or the corresponding transfer functions $H(\omega)$,

as required by ISO 3382 for the determination of room acoustical parameters. All such methods, which may be implemented as FFT, de-convolution or correlation techniques applied to the measurement and processing of the impulse response [1] are able to demonstrate reliable results within normal measurement conditions, i.e. conditions which are linear and time invariant. Any violation of the assumption of linearity and time-invariance will appear as contribution to the existing noise component.

Linearity may be a problem of the sound source if driven with too high a signal level. Effects of nonlinearities are noticeable in room acoustics measurements typically if the dynamic range exceeds 50 dB. For measurements used for high-quality audio processing the dynamic range should be larger than 90 dB. Thus, artefacts caused by nonlinearities would be avoided. The effect and the extent of time variances of the system under test, as caused by temperature changes and corresponding changes of the sound speed, by moving objects, or by changes of wind speed, have been demonstrated by [24].

Even if the impulse response measurement itself is not affected by uncertainties, the post-processing of impulse responses, involving filtering and integration algorithms for parameter estimation, contains various degrees of freedom, with significant effects on the determined results [25,26].

4.2 Sources

Sound sources according to ISO 3382 are supposed to be as omnidirectional as possible. This specification reflects the need for comparable and reproducible results. But it also contains two important shortcomings. Typical sound sources such as dodecahedra are spatially extended objects and combine several transducers in order to produce sufficient sound power. The deviations allowed from an ideal monopole radiation provide only rough approximations. The sound pressure level averaged over 30° may deviate as much as 5 dB from the overall average. This does not exclude larger deviations in specific directions exceeding even 5 dB. As a consequence, a specific early reflection, which may be relevant for clarity or spatial impression, may deviate by several dB when comparing different sources, or when the source orientation is changed, without violating the measurement specifications [27,28].

A second problem related to the specification of omnidirectional sources is that the measurement and the derived parameters do not correspond to the situation which will be present in reality, with natural sound sources such as speakers, singers or musical instruments. The extent to which room acoustical parameters are affected by considering “real-life” source directivities, has been demonstrated by [29] and [30]. One approach to solve this problem is discussed in Sect. 5 below.

4.3 Receivers

While the behavior of omnidirectional receivers is well defined, and corresponding measurement instruments are easily available (high-quality condenser microphones), this is not the case for binaural receivers. Standards for dummy heads [31] specify a certain shape of head and torso but not the pinna as the crucial physiological component, thereby allowing differences in IACC and other metrics based on binaural room impulse responses (BRIRs). The fundamental problem of binaural measurements is the individual characteristics of the human ear and the question, in how far parameters derived from these measurements can be generalized to a larger population. The same problem arises, if BRIRs are used for listening tests, such as those suggested in Sect. 3, because the perceptual correlates of room acoustical properties are confounded with artefacts caused by the non-individual HRTF mismatch.

There are various ways to obtain individual HRTFs by measurement or numerical simulation. In this context, numerical simulations based on photographic imaging or scanning methods as well as direct measurements which allow for individual HRTF acquisition within minutes due to efficient methods for system identification [32–34] have opened up new possibilities to capture the individual characteristics of binaural signals.

The question remains whether receiver directivities beyond monaural and binaural microphones and the figure-of-eight receivers used for the lateral fraction parameters in ISO 3382 can deliver measures which are more highly correlated with the spatial qualities of room acoustical perception than traditional parameters. These would require measurement approaches different from those currently used.

5 New Measurement Approaches

The key to a flexible spatial measurement, taking into account the directivities of source and receiver, could be interfacing the spatial transfer functions based on a decomposition of source and receiver into spherical harmonic components. The approach is a logical extension of the classical Ambisonics approach, with the directional sound field at the listener position in a room captured with four channels (B-format) or with a higher number of channels (Higher Order Ambisonics, HOA) which are typically captured with spherical microphone arrays. The sound source, however, is still included with its actual directivity. In a multiple input multiple output (MIMO) approach, the sound source is implemented as a spherical array as well.

For combining the angular components in azimuth and elevation of sound source and receivers, a set of orthonormal

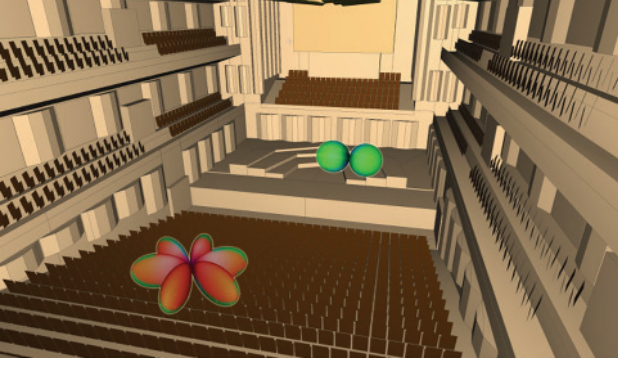


Fig. 1 Measurement (or simulation) of a room transfer function for a particular combination of spherical harmonic components for the source (dipole) and the receiver (hexapole)

basic functions called spherical harmonics (SH) is defined. These are defined as

$$Y_n^m(\vartheta, \varphi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} \cdot P_n^m(\cos \vartheta) \cdot e^{jm\varphi} \quad (3)$$

with n and m denoting the order and degree of the spherical harmonics, respectively. P_n^m is the associated Legendre function whose definition can be found in mathematical textbooks [35, p. 332].

With this basis, any square integrable directional pattern $G(\vartheta, \varphi)$ can be synthesized as a linear combination of SH components Y_n^m , i.e.

$$G(\vartheta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \gamma_{nm} Y_n^m(\vartheta, \varphi) \quad (4)$$

The $Y_n^m(\vartheta, \varphi)$ functions represent monopole, dipole, quadrupole etc. patterns, which must be superimposed in an appropriate way in order to obtain a best match with the specific directional pattern of interest. Similar to the correspondence between time signals and frequency spectra in the Fourier transform, there exists a transformation between the spatial (directivity “balloon”) domain and the SH coefficients. Due to orthogonality, a weighted sum of SH coefficients is a complete and unique representation of the spatial directivity pattern.

In the MIMO technique proposed here, spherical harmonic decompositions are used both for the sound source and for the receiver directivity. Hence, the transfer function for arbitrary sources and receivers can be written as

$$H(\omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \gamma_{nm} \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'} H_{n,m,n',m'}(\omega) \quad (5)$$

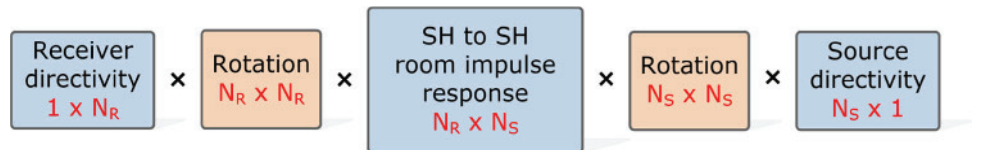
with $H_{n,m,n',m'}(\omega)$ representing the transfer function for a given source $Y_n^m(\vartheta, \varphi)$ and a given receiver $Y_{n'}^{m'}(\vartheta, \varphi)$. If the series in (5) is terminated at $n = N_R$ and $n' = N_S$, $H_{n,m,n',m'}(\omega)$ can be written as a $N_R \times N_S$ matrix of transfer functions for all combinations of SH components on the side of source and receiver. An example for one matrix element for $(n = 1, m = -1, n' = 3, m' = -3)$ is depicted in Fig. 1, where the room acoustical transfer function is simulated for a dipole-to-hexapole configuration.

The challenge is to obtain such complex data by using spherical microphone arrays and spherical loudspeaker arrays at the same time. If, however, this is successful, the acoustic transfer functions can easily be composed post hoc by choosing the appropriate expansion coefficients γ_{nm} , representing the directivities of source and receiver, as weighting coefficients for the measured components $H_{n,m,n',m'}(\omega)$ in (5). Moreover, rotational cues, as they are typically required in dynamic scenes, can easily be obtained by a rotational transformation of $H_{n,m,n',m'}(\omega)$ (Fig. 2).

For the numerical simulation or measurement of the central matrix it is necessary to excite the room with a spherical loudspeaker array and to record with a spherical microphone array. The question which of the two arrays should represent the source and the listener, can be freely selected, because the sound propagation path with its impulse response and transfer function is reciprocal. The decision finally depends on the complexity of the directivity of the source and the receiver, and the best choice is made if the higher-order transducer is used for that part, which requires more spatial resolution.

Spherical microphone arrays, as they are required for the MIMO approach described above, are already available as commercial products. Spherical loudspeaker arrays are, until now, commercially available only as dodecahedron loudspeakers, and can only be driven with all speakers in phase for achieving an omnidirectional radiation. Aiming at loudspeakers for directional radiation, sound sources with adjustable radiation patterns have been designed in an academic environment [36–38]. Hereby, a set of speakers is mounted into a spherical chassis in order to create the radiation of sound sources with a specific directivity pattern. The achievable spatial resolution is limited by the number of loudspeakers used, similar to what we find with limited sampling rates and corresponding aliasing effects. Due to the

Fig. 2 Signal processing chain, including post processing, elements of variable directivities and dynamic rotational transformations at runtime



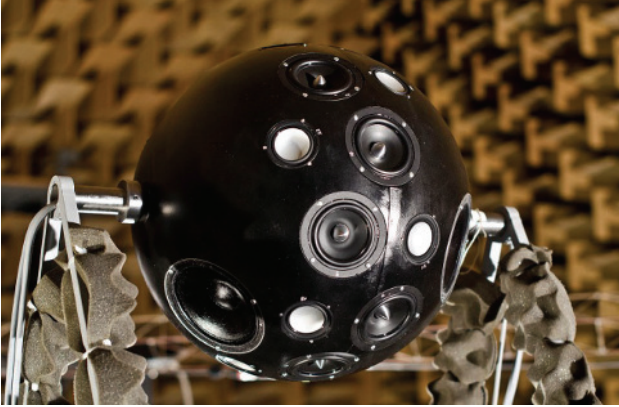


Fig. 3 Spherical loudspeaker source for stepwise sequential measurement

physical size of the single transducers and the requirement for different membrane sizes for a full spectral coverage, the resolution is usually severely limited with these speakers.

Figure 3 shows a recently developed spherical array, used for the measurement and auralization of room impulse responses with an arbitrary directivity pattern of the sound source. With its multi-band excitation, a controlled directivity can be obtained for a frequency range up to 8 kHz including sufficient sound power at low frequencies. Instead of fully covering the sphere with transducers the concept is to use a stepwise rotation in order to cover the whole sphere in a sequential measurement procedure. In the development of the source a simulation model based on a set of analytical descriptions of spherical sound sources was applied, allowing for the analysis of the radiation processes of all individual drivers involved. With this model, an optimum was achieved, considering the transducer size distribution for the various frequency ranges and their distribution on the sphere. Then a SH-based composition was made with an optimization of the positions on the sphere depending on the transducer radius and the space necessary for mounting and separation of the magnets.

This design provides the basis for a synthesis of SH orders up to $N_{\max} = 23$ at about 8 kHz. In first applications, this source served as a Dirac pointer source [39] or as an HRTF source for reciprocal binaural measurements, as presented by [40]. Both studies are based on fast sequential measurements with interleaved sweeps as suggested by [34] and improved by [32].

When applied to room acoustic measurements, distinct early reflections can be evaluated and auralized. This allows for innovative studies on the perception of reflections with regard to their spatial and spectral contribution. The evaluation of scattering surfaces will be of interest as well and possibly new approaches to an in-situ measurement of wall absorption and scattering.

In another study it has been tried to synthesize HRTFs and to apply the source as a binaural receiver in reciprocal measurements. ISO 3382 measurements are usually conducted with omnidirectional sources on the one side, and monopole, dipole or binaural receivers at the other side. Thus, the requirements related to listener directivity are rather high, while the omnidirectionality of dodecahedron loudspeakers is far from being perfect. But with the source array and a 1/2 inch microphone, a quasi-perfect constellation can be found in the reciprocal approach: A perfect omnidirectional microphone transducer on the stage (representing the “source”) and an adaptive SH source (equivalent to a monopole, dipole, or HRTF “receiver”) in the audience. The task remains to transform the SH response in order to obtain the desired equivalent listener directivity, but the procedure still requires just one measurement session in the room, while all data can be obtained in post-processing, including the possibility to construct binaural room impulse responses for individual listeners.

Similar approaches were already applied for the measurement of stage acoustical properties, with the additional challenge that source and receiver might be co-located [41].

6 Conclusions

In room acoustics, we look back on more than 50 years of research on developing psychological measuring instruments for the concept of ‘room acoustical impression’, and on more than 100 years of research on the development of physical measures which could serve as technical predictors for these perceptual qualities. On both sides, the state of the art is surprisingly unsatisfactory. The path towards an improvement of the situation, according to the authors, could lead through applying modern approaches of experimental psychology and test theory on a significantly augmented pool of stimuli, generated by state-of-the-art technologies for room acoustical simulation and auralization. The fundamental perceptual components delivered by this approach will, most likely, not be predictable by traditional room acoustical parameters, but require advanced measurement techniques based on spherical arrays of transducers for both source and receiver characteristics, as well as new auditory models for feature extraction. The physical and psychological aspects of the problem are, in any case, inextricably linked with each other.

Acknowledgments For valuable advice on issues of test design and statistical analysis the authors would like to thank Dr. Steffen Lepa, Audio Communication Group, TU Berlin.

References

1. Müller, S., Massarani, P.: Transfer-function measurement with sweeps. *J. Audio Eng. Soc.* **49**(6), 443–471 (2001)

2. Meyer, J.: *Acoustics and the Performance of Music*. Springer, Berlin (2009)
3. Bregman, A.S.: *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge (1994)
4. Thiering, M.: *Spatial Semiotics and Spatial Mental Models. Figure-Ground Asymmetries in Language*, de Gruyter Mouton (2014)
5. Van Dorp Schuitman, J., de Vries, D., Lindau, A.: Deriving content-specific measures of room acoustic perception using a binaural, nonlinear auditory model. *J. Acoust. Soc. Am.* **133**(3), 1572–1585 (2013)
6. Sabine, W.C.: The accuracy of musical taste in regard to architectural acoustics. In: *Proceedings of the American Academy of Arts and Sciences* 42, 2 (1906), pp. 53–58. Reprinted. In: Sabine, W.C.: *Collected Papers on Acoustics*, 71–77. Harvard University Press, Cambridge (1922)
7. Bagenal, H.: Designing for musical tone. *J. R. Inst. Br. Archit.* **32**(20), 625–629 (1925)
8. Knudsen, V.O.: Acoustics of music rooms. *J. Acoust. Soc. Am.* **2**, 434–467 (1931)
9. Sabine, P.E.: The acoustics of sound recording rooms. *Trans. Soc. Motion Pict. Eng.* **12**(35), 809–822 (1928)
10. Watson, F.R.: *Acoustics of Buildings*. Wiley, New York (1923)
11. ABS 300: Department of Commerce. Bureau of Standards (ed.): *Architectural Acoustics. Circular of the Bureau of Standards, No. 300*, Washington 1926. Cf. <http://www.babel.hathitrust.org/cgi/pt?id=uiug.30112106642967>
12. Somerville, T., Gilford, C.L.S.: Acoustics of large orchestral studios and concert halls. In: *The Proceedings of the Institution of Electrical Engineers* 104, pp. 85–97 (1957). Reprinted. In: *J. Audio Eng. Soc.* **7**(4), 160–172 (1959)
13. Beranek, L.L.: *Music, Acoustics and Architecture*. Wiley, New York (1962)
14. Hawkes, R.J., Douglas, H.: Subjective acoustic experience in concert auditoria. *Acustica* **24**(5), 235–250 (1971)
15. Lehmann, P., Wilkens, H.: Zusammenhang subjektiver Beurteilungen von Konzertsälen mit raumakustischen Kriterien. *Acustica* **45**, 256–268 (1980)
16. Wilkens, H.: Mehrdimensionale Beschreibung subjektiver Beurteilungen der Akustik von Konzertsälen. *Acustica* **38**, 10–23 (1977)
17. Berg, J., Rumsey, F.: Identification of quality attributes of spatial audio by repertory grid technique. *J. Audio Eng. Soc.* **54**(5), 365–379 (2006)
18. Lokki, T., Pätynen, Y., Kuusinen, A., Tervo, S.: Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. *J. Acoust. Soc. Am.* **132**(5), 3148–3161 (2012)
19. Lindau, A., Erbes, V., Lepa, S., Maempel, H.-J., Brinkman, F., Weinzierl, S.: A spatial audio quality inventory (SAQI). *Acta Acust. United Acust.* **100**, 984–994 (2014)
20. Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J.: Evaluating the use of explanatory factor analysis in psychological research. *Psychol. Methods* **4**, 272–299 (1999)
21. Gade, A.C.: Acoustics for symphony orchestras; status after three decades of experimental research. In: *Proceedings of the International Symposium on Room Acoustics (ISRA 2010)*, pp. 1–12. Melbourne, Australia (1999)
22. ISO 3382–2: *Measurement of Room Acoustic Parameters Part 2: Reverberation Time in Ordinary Rooms* (2008)
23. ISO 3382–1: *Measurement of Room Acoustic Parameters Part 1: Performance Rooms* (2009)
24. Vorländer, M., Kob, M.: Practical aspects of MLS measurements in building acoustics. *Appl. Acoust.* **52**, 239–258 (1997)
25. Guski, M., Vorländer, M.: Comparison of noise compensation methods for room acoustic impulse response evaluations. *Acta Acust. United Acust.* **100**, 320–327 (2014)
26. Katz, B.: International round robin on room acoustical impulse response analysis software. *Acoust. Res. Lett. Online* **5**(4), 158–164 (2004)
27. Leishman, T.W., Rollins, S., Smith, H.M.: An experimental evaluation of regular polyhedron loudspeakers as omnidirectional sources of sound. *J. Acoust. Soc. Am.* **120**(3), 1411–1422 (2006)
28. San Martín, R., Witew, I., Arana, M., Vorländer, M.: Influence of the source orientation on the measurement of acoustic parameters. *Acta Acust. United Acust.* **93**, 387–397 (2007)
29. Otondo, F., Rindel, J.H.: The influence of the directivity of musical instruments in a room. *Acta Acust. United Acust.* **90**, 1178–1184 (2004)
30. Wang, L.M., Vigeant, M.C.: Evaluations of output from room acoustic computer modeling and auralization due to different sound source directionalities. *Appl. Acoust.* **69**(12), 1281–1293 (2008)
31. SS-IEC 959: *Acoustics—Provisional head and torso simulator for acoustic measurements on air conduction hearing aids* (1991)
32. Dietrich, P., Masiero, B., Vorländer, M.: On the optimization of the multiple exponential sweep method. *J. Audio Eng. Soc.* **61**(3), 113–124 (2013)
33. Fallahi, M., Brinkmann, F., Weinzierl, S.: Simulation and analysis of measurement techniques for the fast acquisition of head-related transfer functions. In: *Proceedings of German Annual Conference on Acoustics (DAGA)*, Nuremberg 2015, in print (2015)
34. Majdak, P., Balazs, P., Laback, B.: Multiple exponential sweep method for fast measurement of head-related transfer functions. *J. Audio Eng. Soc.* **55**(7/8), 623–637 (2007)
35. Abramowitz, M., Stegun, I.A.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York (1965)
36. Behler, G.: *Sound source with adjustable directivity*. ASA Autumn meeting, Honolulu (2006)
37. Pollow, M., Dietrich, P., Kunkemöller, M., Vorländer, M.: Synthesis of room impulse responses for arbitrary source directivities using spherical harmonics decomposition. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY (2011)
38. Schwenke, R., Zotter, F., Schmeder, A., Wessel, D.: Room acoustics measurements with an approximately spherical source of 120 drivers. In: *Proceedings of Acoustics 08 Paris* (2008)
39. Pollow, M., Klein, J., Dietrich, P., Behler, G., Vorländer, M.: Optimized spherical sound source for room reflection analysis. In: *Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENC 2012)*, Aachen (2012)
40. Pollow, M., Klein, J., Dietrich, P., Vorländer, M.: Including directivity patterns in room acoustical measurements. In: *Proceedings ICA Montreal, Canada* (2013)
41. Miranda, L., Cabrera, D., Stewart, K.: A concentric compact spherical microphone and loudspeaker array for acoustical measurements. In: *135th Audio Engineering Society Convention*, New York, USA (2013)